

## **Informatique et histoire**

### **Le codage automatique des professions et des provenances**

Cl. DESAMA

Pourquoi recourir à des procédures automatiques pour coder les informations qui doivent l'être? La raison en est double. Tout d'abord assurer au codage un maximum de rigueur et de précision. Ensuite veiller à ce que cette opération ne retarde pas le déroulement du travail. Il est évident en effet que le codage manuel des professions et des provenances dans un fichier de plusieurs milliers d'articles constitue une tâche lourde et lente que l'utilisateur doit assumer lui-même pour la plus grande part. Le recours à l'ordinateur représente à ce point de vue un progrès considérable.

Le caractère répétitif de certains codages (même profession = même code) limite en effet l'intervention de l'utilisateur à l'attribution d'un code aux provenances ou professions nouvelles ou encore aux variantes orthographiques de celles qui sont déjà codées. Ainsi pouvons-nous bénéficier dans une très large mesure de la vitesse d'exécution de la machine.

Sur le plan de la rigueur et de la précision l'ordinateur apporte également un avantage décisif. Quel que soit le soin que l'on puisse apporter au codage manuel, celui-ci n'est jamais exempt d'erreur soit par l'attribution abusive d'un code soit par distraction pure et simple. Il faut donc prévoir une ou plusieurs relectures du fichier codé et opérer les corrections qui s'imposent. Plus le fichier est grand, plus cette opération est hasardeuse et plus la probabilité de laisser subsister des erreurs est grande. Le codage automatique présente l'avantage

de limiter ces risques à une partie restreinte du fichier. Par ailleurs, la probabilité de voir l'ordinateur attribuer un code différent à deux professions identiques est nul.

Si le codage automatique se révèle à l'expérience plus rapide, plus efficace et plus précis que le codage manuel, il exige néanmoins que soient remplies au préalable certaines conditions:

1. Disposer d'un code de bonne qualité, bien adapté à la documentation à laquelle on l'applique. Si l'on veut tirer le profit maximal du recours à l'ordinateur, il faut que le plus grand nombre possible de professions aient d'emblée un code.
2. Organiser les fichiers de telle manière que les codes attribués manuellement viennent enrichir le code initial. Ainsi au fil des codages, le résidu se ramènera peu à peu aux seules professions nouvelles.
3. Mettre au point des procédures suffisamment souples pour que les professions ou les provenances codées manuellement soient réintroduites dans les fichiers sans entraîner des réécritures ou des traitements supplémentaires toujours coûteux.

## 1. Le codage socio-professionnel

Le principe du codage automatique est le suivant: à partir d'un code socio-professionnel préenregistré sous la forme d'un thésaurus, l'ordinateur attribue à chaque profession son code correspondant. Au terme de l'opération il reste un résidu comprenant les mentions professionnelles ou les variantes qui ne figurent pas dans le thésaurus. Voyons en détail les différentes phases du processus.

### 1.1. L'enregistrement du code socio-professionnel

Le fichier du code socio-professionnel comporte 900 articles auxquels correspondent autant d'enregistrements. Sur chaque *record* figure la mention "en clair" de la profession ou du statut ainsi que le code qui lui est affecté. La structure d'enregistrement se présente comme suit avec un maximum de 80 positions :

5 7 0 1 3 1	C O R R O Y E U R
numéro de code	mention socio-professionnelle

Le numéro de code occupe six positions dont les quatre premières sont réservées à la profession et les deux dernières au statut. Quant à la mention socio-professionnelle, elle est encodée puis enregistrée telle qu'elle figure dans la liste nominative de recensement. Elle s'étend donc sur une longueur variable dans une zone bornée par les positions 8 et 80.

Avant même les traitements de codage on imprime une première édition du code d'abord suivant l'ordre croissant des numéros par secteur économique, par profession et par statut, ensuite suivant l'ordre alphabétique des professions. Chaque liste obtenue est soigneusement référenciée car elle indique l'état du code à un moment déterminé.

Une fois trié selon l'ordre alphabétique des professions, le fichier est prêt pour la phase suivante: l'indexation séquentielle. Nous ne reviendrons pas sur cette technique et sur les conditions nécessaires à son utilisation; la succession des opérations est indiquée à la figure 1.

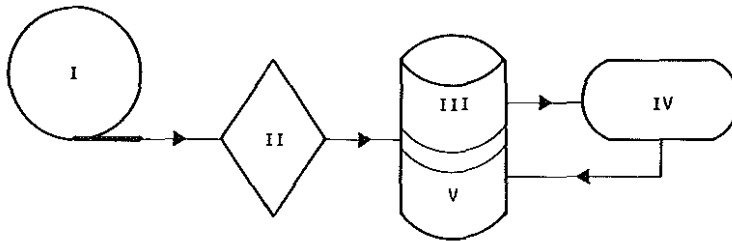


fig. 1  
L'indexation séquentielle

- I Code socio-professionnel enregistré sur bande
- II Tri du fichier ainsi enregistré suivant l'ordre alphabétique des mentions socio-professionnelles
- III Ecriture sur disques du fichier trié
- IV Indexation séquentielle du fichier trié
- V Réécriture sur disques du fichier indexé

C'est ce fichier indexé qui va servir de thésaurus dans l'opération proprement dite de codage.

## 1.2. Le codage

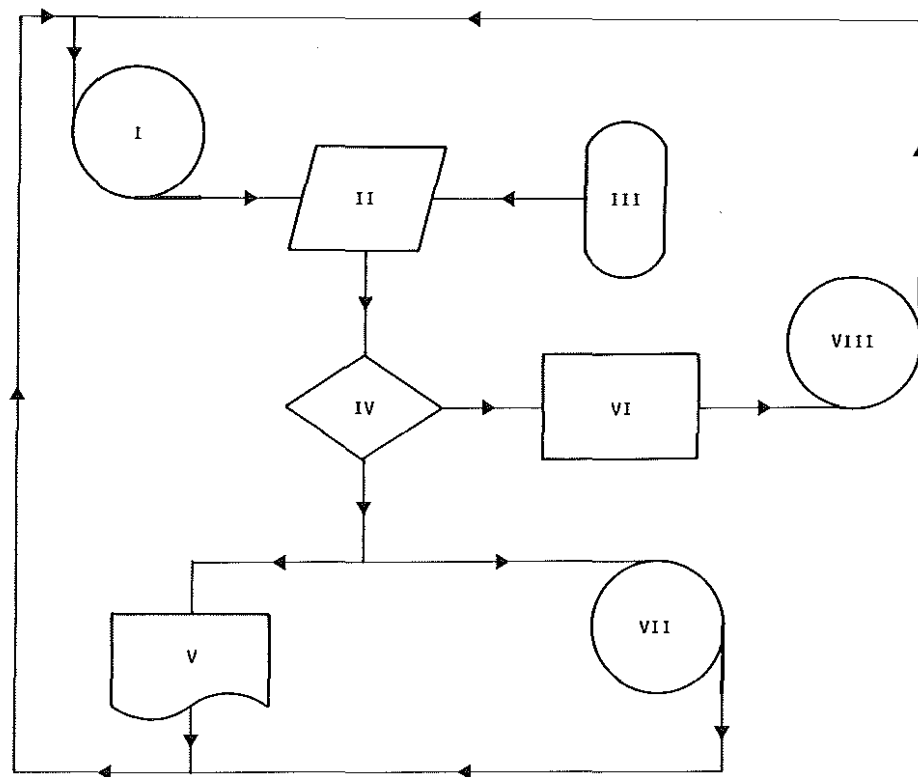


fig. 2  
Le codage automatique

- I Fichier à coder
- II Lecture des deux fichiers
- III Fichier "code socio-professionnel" enregistré en indexé séquentiel
- IV Question : la mention professionnelle lue dans le fichier à coder trouve-t-elle son pendant exact dans le fichier du code socio-professionnel ? Dans l'affirmative l'ordinateur attribue le code voulu en l'écrivant dans une zone laissée vierge du *record* traité du fichier I (VI). Cette zone est, bien entendu, de longueur fixe et couvre les positions 53 à 58. Dans la négative, l'ordinateur imprime la mention professionnelle (V).
- VII Réécriture du fichier codé sur une autre bande ou sur une autre partie de la bande initiale. Il s'agit là d'une opération en soi facultative mais qui offre l'avantage de conserver intact le fichier de départ aussi longtemps qu'on n'est pas certain de la bonne marche des programmes. C'est ce fichier ainsi enregistré qui constitue la base documentaire des traitements ultérieurs.
- VIII Enregistrement sur bande magnétique des *records* dont les mentions professionnelles n'ont pas de correspondant dans le code. L'éclatement du fichier initial en vue de séparer ce qui est codé de ce qui ne l'est pas encore permet, au moment du codage définitif, d'éviter une relecture complète de ce fichier.

Le programme de codage automatique se déroule selon le schéma représenté à la figure 2.

Le codage automatique procède donc par comparaison entre la mention socio-professionnelle recensée et celle qui figure dans le thésaurus. L'indexation séquentielle du code rend cette opération extrêmement rapide et pour départager les professions déjà codées de celles qui ne le sont pas et pour attribuer les codes adéquats aux professions qui figurent dans le thésaurus. Voyons les instructions utilisées à cet effet.

Tout d'abord la fonction ON KEY qui décèle l'absence dans le *record* traité d'une clé d'accès au fichier indexé, en l'occurrence une profession qui figure dans le code. Elle se présente comme suit :

```
ON KEY (SI) GOTO IMPR ;
```

SI étant le nom du fichier indexé et IMPR le label identifiant les instructions relatives à l'impression des professions non codées.

Si ce test est négatif, c'est-à-dire si le *record* contient une clé, on procède au codage et à la réécriture de l'enregistrement codé :

```
(1) READ FILE (VERV 1) INTO (TAB) ;  
(2) READ FILE ( SI ) INTO ( ZL ) KEY (SUBSTR(TAB,80,34)) ;  
(3) SUBSTR (TAB, 53,6) = ZA ;  
(4) REWRITE FILE (VERV 2) FROM (TAB) ;
```

- (1) Lecture du fichier initial (VERV 1) dans la zone TAB qui correspond à la structure d'enregistrement des *records*;
- (2) lecture du fichier indexé (SI) avec comme clé de consultation l'information qui, dans la zone TAB, commence à la position 80 et s'étend sur 34 positions c'est-à-dire, en l'espèce, la profession;
- (3) transfert du code chiffré (ZA) dans une zone de TAB qui comprend les positions 53 à 58 incluse;
- (4) réécriture du *record* codé dans le fichier définitif (VERV 2). Sur base d'un exemple concret, le déroulement de l'opération est représenté à la figure 3.

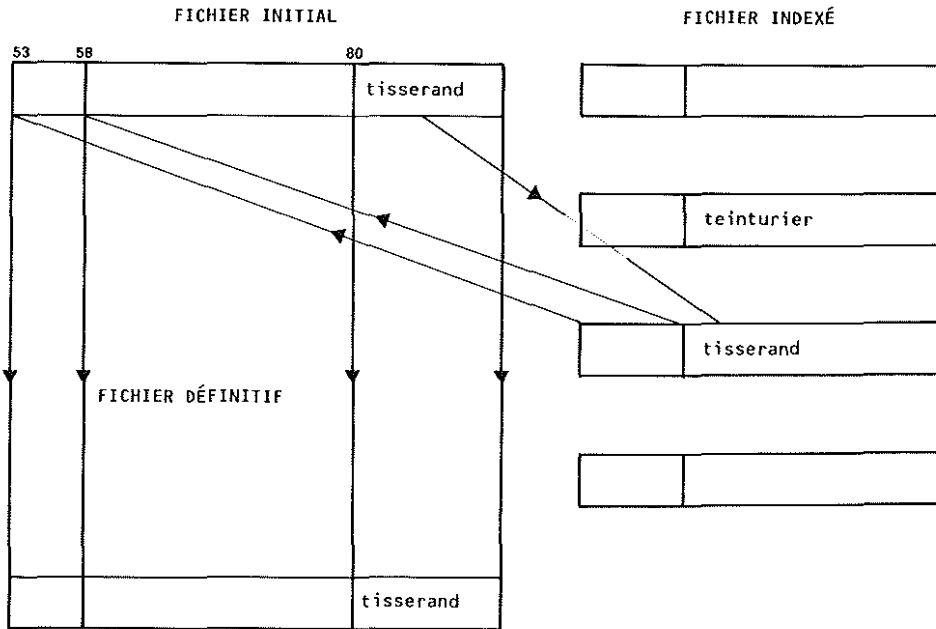


fig. 3  
L'opération de codage

### 1.3. Le codage définitif

Nous entendons par là l'ultime phase du codage qui concerne les enregistrements dont les mentions professionnelles ne figuraient pas dans le code. Détectés par la fonction ON KEY, ces professions ou statuts sociaux sont imprimés par l'ordinateur au fur et à mesure de l'apparition et perforés automatiquement sur cartes suivant le modèle que nous avons décrit dans le paragraphe précédent. Sur cette liste ainsi dressée, nous trouvons non seulement des mentions absentes du code mais encore des variantes graphiques. La forme *tisserant* sera considérée par l'ordinateur comme une profession différente de *tisserand* et par là-même non codée. Même si cette forme erronée se présente à plusieurs reprises dans le fichier, elle n'est retenue qu'une seule fois par l'ordinateur en vue de l'impression. Ainsi le résidu du codage ne comprendra-t-il que des mentions nouvelles ou erronées *différentes*. C'est donc un ensemble de professions ou de statuts d'importance réduite que l'utilisateur devra coder manuellement.

Une fois les codes attribués, comment s'effectue le codage de l'ensemble des formes résiduelles du fichier? Schématiquement l'enchaînement des opérations est représenté à la figure 4.

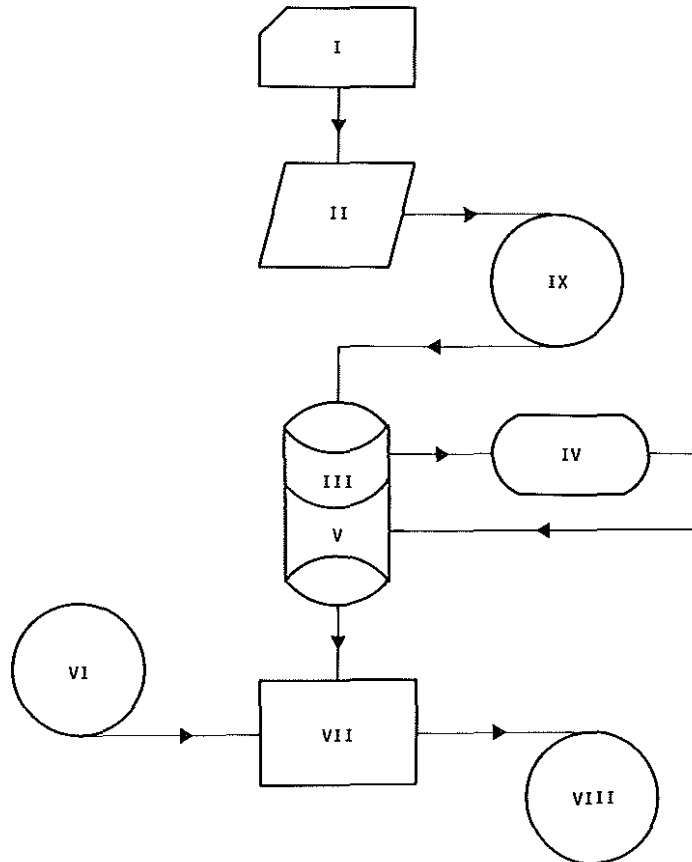


fig. 4  
Le codage des formes résiduelles

Les enregistrements avec les mentions professionnelles non codées sont complétés par le code adéquat (I) puis lues par l'ordinateur (II). Elles sont ensuite enregistrées, en juxtaposition (MOD, PASS) sur la bande magnétique qui contient le code socio-professionnel (IX). Le nouveau fichier "code" ainsi constitué est alors trié suivant l'ordre alphabétique des mentions

professionnelles (III) avant d'être indexé séquentiellement puis recopié sur les disques (V). Il ne reste plus dès lors à l'ordinateur qu'à coder (VII) les *records* qui ne l'étaient pas (fichier VI) et à écrire (VII) les enregistrements complets sur la bande où figurent les *records* déjà codés (VIII).

Nous disposons alors d'un fichier définitif homogène où toutes les informations recensées au titre de la profession se trouvent tant dans leur graphie originale que sous leur forme symbolique.

Sur le plan de la programmation, la procédure fait appel à deux reprises à une instruction OS (Operating System) qui permet de juxtaposer, par enregistrement, un fichier à un autre sur une même bande ou un même segment de bande en manière telle que l'ordinateur puisse les lire et, éventuellement, les trier comme s'il s'agissait d'un seul fichier. Cette instruction intervient au niveau de la définition de la DISP. Elle s'exprime de la façon suivante :

```
DISP = ( MOD, PASS )
```

l'attribut MOD commandant cette juxtaposition de fichiers.

Une pareille possibilité nous est extrêmement utile en matière de codage automatique car elle nous évite de recommencer l'ensemble de la procédure à chaque étape et, par-là même, de manipuler des fichiers importants. Elle ne contribue pas peu à faire du codage automatique un traitement souple et aisément contrôlable.

Sur le plan de la programmation, l'indexation séquentielle du fichier s'effectue par le biais d'instructions relativement simples qui agissent à trois niveaux :

- dans les déclarations, le fichier à indexer est défini de la manière suivante :

```
DCL CODE FILE RECORD OUTPUT KEYED  
ENV (INDEXED KEYLENGTH (4) KEYLOC (2)) ;
```

*Explications:* Il s'agit d'un fichier de sortie (OUTPUT) que l'on consulte au moyen d'une clé (KEYED). Les clés sont indexées (INDEXED), elles occupent quatre positions (KEYLENGTH) et commencent à la position 2 de l'enregistrement (KEYLOC);

- dans les commandes d'écriture :

```
WRITE FILE (CODE) FROM (TAB) KEYFROM (SUBSTR(TAB,2,4)) ;
```



*Explications* : on écrit sur disque le fichier CODE dans une structure d'enregistrement définie par TAB avec comme clé l'information qui se trouve dans TAB entre les positions 2 et 5 incluses;

- dans les instructions systèmes OS relatives au fichier CODE et plus précisément dans le DCB :

```
DCB = (RECFM=VB, LRECL=364, BLKSIZE=3644, DSORG=IS)
```

*Explications* : la DCB définit au niveau du système le format du *record*, en l'occurrence il s'agit d'un enregistrement en longueur variable bloquée (VB), sa longueur (LRECL) en nombre de *bytes* ou positions (364) et la taille du fichier (BLKSIZE). L'attribut DSORG = IS est propre au fichier indexé séquentiel; il commande à l'ordinateur une organisation de fichier qui permet l'indexation. Il est à noter toutefois que quatre positions doivent être ajoutées au BLOCKSIZE en raison du format VB des *records*.

## 2. Le codage des provenances

La conception générale du codage automatique des provenances est analogue à celle que nous avons décrite à propos du codage des professions et des statuts sociaux. Nous ne nous attarderons pas dès lors à souligner les similitudes et à répéter ce que nous avons déjà expliqué précédemment à propos des procédures et des techniques de correction (UPDATE) ou d'indexation séquentielle. Examinons plutôt les singularités qui donnent au codage des provenances son caractère spécifique.

### 2.1. L'enregistrement du code provenance

Le fichier du code provenance contient autant d'articles qu'il y a de communes belges et d'entités administratives ou géographiques retenues pour l'étranger. L'information relative aux communes se répartit en quinze rubriques tantôt de longueur fixe, tantôt de longueur variable. Chaque zone de longueur variable est séparée de celle qui la suit par un astérisque dont le rang dans la notice définit l'objet de la rubrique.

Une exception a cependant été consentie à ces principes d'enregistrement. La dernière partie du *record* est en effet introduite par le signe + et contient des lettres (A,B,C,D, ...) comme séparateurs de rubriques. Si l'on veut comprendre cette disposition particulière, il faut rappeler que la dernière zone de la notice a une vocation bien précise. Elle est destinée à recevoir des renseignements tels que le chiffre de la population, la nature socio-économique de la commune, etc. Nous avons introduit des signes différents de manière à pouvoir extraire, si nécessaire, ces informations de l'enregistrement et constituer un fichier annexe qui fera l'objet, éventuellement, d'un traitement séparé. Le schéma d'enregistrement d'une notice complète est représenté à la figure 5.

```

- 2371 2555 1436 63083 4800      B      * VERVIERS *
   1    2    3    4    5    6    7    8    9

LIEGE * VERVIERS * VERVIERS * VERVIERS * VERVIERS *
 10    11    12    13    14

4271119300722 + 2371 A008718 B008725 C009974 D009881
      15          16    17    18    19    20

E009965 F012335 G015832
  21    22    23

```

fig. 5

## Schéma d'enregistrement d'une notice complète

- 1 code commune
- 2 coordonnée hectométrique (abscisse)
- 3 coordonnée hectométrique (ordonnée)
- 4 code INS
- 5 code postal
- 6 code de la nouvelle commune (blanc dans le cas de Verviers)
- 7 lettre distinctive du pays
- 8 autres codes de provenance éventuels (blanc dans le cas de Verviers)
- 9 nom de la commune
- 10 nom de la province
- 11 nom de l'arrondissement judiciaire
- 12 nom du canton judiciaire (1947)
- 13 nom du canton judiciaire (1971)
- 14 nom de l'arrondissement administratif
- 15 code indiquant la scission, l'annexion ou le regroupement avec d'autres communes
- 16 début des données provenant de la dernière carte avec rappel du code de la commune
- 17 )
- 18 )
- 19 )
- 20 ) chiffres de la population aux recensements
- 21 )
- 22 )
- 23 )

```

1806
1808
1810
1812
1816
1828
1846

```

Une fois enregistré sur bande, le fichier des provenances est appelé à jouer un rôle analogue au fichier socio-professionnel, c'est-à-dire celui d'un thésaurus dans lequel à chaque mention de lieu correspond un code chiffré. Il suscite cependant quelques problèmes qu'il convient de résoudre avant d'entreprendre l'indexation. Tout d'abord les amphibologies. Ainsi la provenance "Villers" qui figure dans un enregistrement peut être interprétée de 21 façons différentes depuis Villers-aux-Tours jusqu'à Villers-sur-Semois en passant par Villers-le-Temple. Quelle mention l'ordinateur va-t-il choisir? Quel code affectera-t-il? Faute d'instructions précises l'ordinateur n'effectuera aucun choix, ne prendra aucune décision et la procédure de codage s'en trouvera ainsi interrompue. Ce type de difficultés doit donc être résolu au préalable mais comment? Sauf à posséder les éléments nécessaires à une identification irréfutable, nous nous voyons contraints d'avoir recours à ce que nous appelons un code d'indétermination, en l'occurrence 0000. Ce code est attribué par l'ordinateur chaque fois qu'il rencontre une mention non exclusive d'un lieu. Lorsque la fréquence des provenances indéterminées reste faible, c'est-à-dire non significative sur le plan statistique, il n'est pas nécessaire de lever l'indétermination. En revanche si le seuil de 5 % est dépassé, nous répartissons les provenances indéterminées entre les différents codes possibles proportionnellement à leur nombre relatif d'apparitions dans le fichier. Prenons l'exemple de Marche qui est présent, sous cette forme abrégée, dans vingt enregistrements. Pour lever l'indétermination, nous devons choisir entre deux possibilités : Marche-en-Famenne (1404) et Marche-les-Dames (1405). La première de ces mentions apparaît huit fois et la seconde deux fois: soit 80 % pour l'une et 20 % pour l'autre. Dès lors les vingt enregistrements qui contiennent la provenance Marche vont être codés comme suit : les seize premiers recevront le code 1404, les quatre derniers le code 1405.

Les cas d'homonymie sont beaucoup plus malaisés à résoudre. Une règle simple consisterait à attribuer systématiquement le code de celle des deux communes qui est le plus proche du point d'arrivée. Elle se heurte pourtant à deux objections : l'une à caractère statistique et l'autre fondée sur l'argument de vraisemblance. Sur le plan statistique, l'application de cette règle à une homonymie de fréquence significative aurait pour conséquence d'introduire un biais systématique. Par ailleurs nous ne pourrions retenir comme vraisemblable l'hypothèse selon laquelle, par exemple, les habitants originaires de Mons proviendraient de la commune de la banlieue liégeoise plutôt que du

chef-lieu de la province de Hainaut. La règle de la plus grande proximité doit donc être maniée avec beaucoup de prudence et, dans certains cas, il est préférable de conserver l'indétermination ou, provisoirement, de laisser au chercheur lui-même le soin de prendre une décision *a posteriori* c'est-à-dire après le codage automatique.

Les homonymies entre noms de lieux créent une autre difficulté au moment de l'indexation du fichier. En effet, pour indexer un fichier, les clés doivent être obligatoirement différentes sous peine de rendre impossible sa consultation. Dans la mesure où

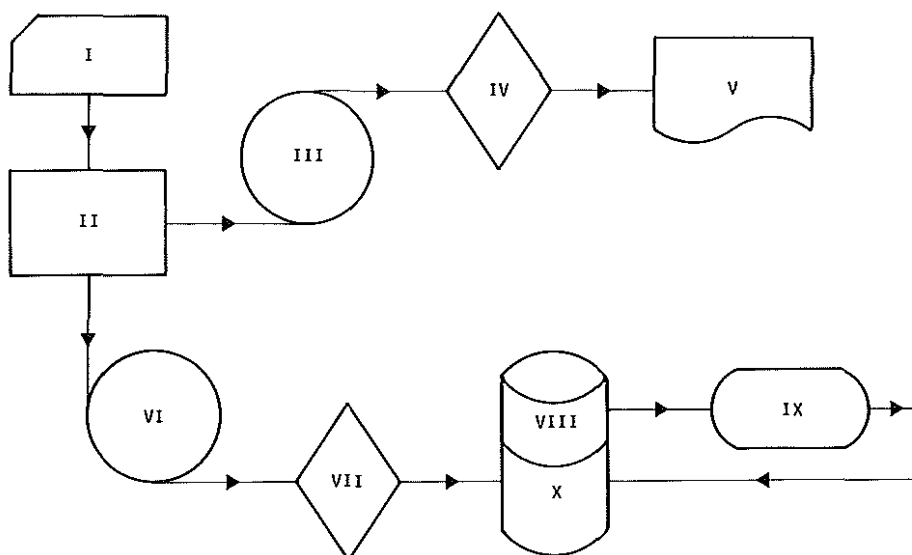


fig. 6  
La préparation du codage des provenances

- I Fichier des codes de provenances
- II Sélection des provenances homonymes
- III Enregistrement sur bande des provenances homonymes
- IV Tri alphabétique des provenances homonymes
- V Liste des provenances homonymes
- VI Enregistrement sur bande du code provenance épuré des homonymes
- VII Tri alphabétique du fichier "code-provenance"
- VIII Ecriture sur disques du fichier trié
- IX Indexation séquentielle du fichier
- X InSCRIPTION sur disques du fichier indexé

la mention de provenance est utilisée comme clé, les homonymes (par ex. Mons) sont nécessairement exclus du fichier que l'on désire indexer de façon séquentielle.

Que ce soit pour des raisons touchant à l'exactitude des résultats ou pour des motifs d'ordre technique, on constate qu'il est indispensable de traiter séparément les homonymes, en dehors des procédures de codage automatique. C'est au niveau de l'enregistrement du fichier-code qu'il convient donc de sélectionner les provenances homonymes et de constituer un fichier annexe susceptible d'inspections approfondies. Compte tenu de ces éléments, le schéma des opérations préalables au codage est représenté à la figure 6.

## 2.2. Le codage

Le programme de codage automatique des provenances se déroule selon un schéma analogue à celui des professions. Mêmes principes de traitement et, partant, mêmes instructions. Le code est écrit dans une zone vierge de l'enregistrement après consultation du fichier des provenances codées. La zone utilisée est de format fixe et occupe les positions 60 à 63 incluse. Lorsque la mention de provenance ne trouve pas son correspondant dans le code, elle est imprimée, ce qui nous permet de disposer, comme dans le cas de professions, d'un document aisé à utiliser en vue du codage manuel.

Les mentions qui sont ainsi rejetées par l'ordinateur constituent, dans la plupart des cas, des variantes orthographiques de noms de communes identifiées. Les mentions nouvelles se répartissent entre, d'une part, des noms de localités ou de régions de pays étrangers et, d'autre part, des lieux-dits ou des hameaux qui se trouvent sur le territoire d'une commune belge. Cette dernière catégorie est la plus malaisée à coder car il n'est pas toujours possible de les identifier avec succès ou de les attribuer à une commune plutôt qu'à une autre.

Si le codage des professions ne permettait pas de résoudre entièrement toutes les occurrences, le résidu de mentions échappant à la codification demeurerait néanmoins fort limité (entre 0,2 et 0,5 %). La situation est moins favorable dans le cas des provenances puisqu'aux lieux non identifiés viennent s'ajouter les indéterminations. La proportion de mentions non

codables est donc plus élevée, elle s'inscrit entre 3 % et 4 %, ce qui n'affecte pas d'une manière significative les résultats ultérieurs.

### 2.3. *Le codage définitif*

Une fois l'identification de la provenance menée à bien, on est en mesure de lui attribuer le code adéquat. Les enregistrements ainsi complétés sont ensuite réécrits sur la bande où figure le code-référence en vue de la mise à jour permanente du thésaurus des provenances codées.

### 2.4. *L'utilisation du code provenance*

Contrairement au code socio-professionnel, le code provenance ne constitue pas en soi une source d'information. Il n'est qu'un intermédiaire entre l'article (*record*) du fichier où figurent les données à exploiter et le fichier de référence où sont inscrites l'ensemble des informations afférentes à chacune des provenances.

Les quatre chiffres du code forment la clé de consultation du fichier des communes et des noms de lieux, préalablement organisé selon la technique de l'indexé-séquentiel.

Toute exploitation des données de provenance par l'ordinateur suppose donc que soient réunies trois conditions :

- a) l'affectation d'un code à chaque provenance;
- b) l'indexation séquentielle du fichier de référence (fichier des communes et des noms de lieux) selon le critère du code;
- c) le choix des variables à isoler dans ce fichier de références.

Dans les pages qui précèdent nous avons décrit les différentes phases du codage automatique qui visent précisément à attribuer un code chiffré à chaque provenance. Moyennant les restrictions que nous avons formulées précédemment, on peut considérer qu'au terme de cette opération la première condition est réalisée. Restent l'indexation séquentielle du fichier et le choix des variables.

#### 2.4.1. *L'indexation séquentielle du fichier*

Elle s'effectue en deux temps dans un ordre obligatoire :

- le tri du fichier avec comme critère la donnée qui constitue la clé de consultation;
- l'indexation proprement dite commandée par des instructions de programme au niveau de la procédure et au niveau du système.

Pour pouvoir réaliser l'indexation séquentielle, l'ordinateur doit disposer impérativement d'un fichier trié selon l'ordre croissant des clés, de manière telle qu'aucune clé ne s'y trouve à plus d'une reprise. Il est donc impératif de recourir au fichier des lieux et des provenances dans son état initial où, à un code, ne correspond qu'une seule information. Rappelons, en effet, que l'enrichissement du fichier, indispensable pour améliorer le rendement du codage automatique, implique que les différentes variantes d'une même provenance se trouvent reprises comme des informations indépendantes mais affectées du même code. Cette version du fichier est, par conséquent, inutilisable au stade de l'exploitation.

#### 2.4.2. *Le choix de variables*

Le fichier des communes belges se présente davantage comme un dictionnaire aux multiples fonctions que comme un ensemble immédiatement opérationnel. Il faut donc choisir puis mettre en évidence, au niveau de l'enregistrement, celles des données qui vont faire l'objet du traitement.

En ce qui concerne les provenances, nous avons retenu les variables suivantes : nom de la commune, coordonnées hectorométriques, division administrative, chiffre de population à l'époque considérée.

Le choix de ces variables ne s'opère pas arbitrairement. Il est fonction de la nature des résultats que l'on désire obtenir et peut changer d'un travail à l'autre. Pour notre part nous désirons mesurer les flux migratoires à l'aune des distances (coordonnées hectorométriques), en relation avec un cadre administratif qui, s'il n'a pas toujours une signification économique et sociale, n'en constitue pas moins un élément de comparaison dans le temps. La commune, l'arrondissement et la province définissent en effet autant de catégories statistiques qui

figurent dans les documents officiels jusqu'à nos jours. Quant au chiffre de la population, il permet d'interpréter l'ampleur des mouvements migratoires en fonction de l'importance démographique du lieu de départ.

Ces différentes variables sont appelées par ailleurs à être combinées avec d'autres données du recensement et, en particulier, avec l'âge, la profession ou le statut social.

L'avantage d'un fichier, tel que celui des communes belges, organisé sous forme de banque de données est précisément d'offrir à l'utilisateur la possibilité de choisir ses paramètres et de modifier éventuellement son choix en cours de traitement.